

1 GC Cgt Cactcc Ctcgt Cat Cgataacat Cct Gtc CAA GAT Cg A GAA Cg A Gt A Cg
 1> Al a Val Thr Pro Ser Val Ile Asp Asn Ile Leu Ser Lys Ile Gl u Asn Gl u Tyr G
 56 AGGT GCT Gt ACC TGA Gcc Gt Gcagg Ggt Cta CC GAG CC TGA GAA GcAG
 19> Iu Val Leu Tyr Leu Lys Pro Leu Al a Gl y Val Tyr Arg Ser Leu Lys Lys Gl n
 109 Ct GgA GAA CaACGt GAT GAC Ctt CAA CGT Gt A CGT GAGGATAT CT GAA CAGC
 37> Leu Gl u Asn Asn Val Met Thr Phe Asn Val Asn Val Lys Asp Ile Leu Asn Ser
 163 CG GTT CAA CAA GOG GGA GAA CTT CAA GAA CGT CC TGA GAGC GAT Ct GAT Ccc
 55> Arg Phe Asn Lys Arg Gl u Asn Phe Lys Asn Val Leu Gl u Ser Asp Leu Ile Pr
 216 Cta CAA GgAT Ct Gac CAG CAG CAA Cta CGt Ggt CAA GGATCC Cta CAA GTT CC
 72> o Tyr Lys Asp Leu Thr Ser Ser Asn Tyr Val Val Lys Asp Pro Tyr Lys Phe L
 269 t GAA CAA GgA GAA GAGAGATAA Gt Ct GAGCAGTT CAA Cta Cat CaAGGATAG
 90> eu Asn Lys Gl u Lys Arg Asp Lys Phe Leu Ser Ser Tyr Asn Tyr II e Lys Asp Se
 324 Cat Tgatac Cgat AT CAA Ctt Cgc CAA CGAT GT CCT GggATA Cta CAA GAT CC
 108> r II e Asp Thr Asp Ile Asn Phe Al a Asn Asp Val Leu Gl y Tyr Tyr Lys II e Le
 378 Gt CCGA GAA Gt A CAA GAGC GAT Ct GgAT TCA AT CAA GAA Gt A CAT CAAC GA TAA
 126> u Ser Gl u Lys Tyr Lys Ser Asp Leu Asp Ser II e Lys Lys Tyr II e Asn Asp Ly
 432 Gc A Ggg Aga Gaa Cgagaa Gt ac Ct Gcc TT CC TGA CAA CAt CgAG ACC Ct Gt A
 144> s Gl n Gl y Gl u Asn Gl u Lys Tyr Leu Pro Phe Leu Asn Asn II e Gl u Thr Leu Ty
 486 CAA Gac Cgt Caa Cgataa GATTGAT Ct Gtt Cgt Gt Cca CC TGA Gcc Caa Ggt
 162> r Lys Thr Val Asn Asp Lys II e Asp Leu Phe Val II e His Leu Gl u Al a Lys Va
 Ndel
 540 Cct GAA Cta Cac AT AT GAG AA GAG CAA CGT GGA Ggt CAA GAT CAA GgA Gct GAA
 180> I Leu Asn Tyr Thr Tyr Gl u Lys Ser Asn Val Gl u Val Lys II e Lys Gl u Leu As
 594 TTAC Ct GAA Gac Cat Cca GgA TAA CC TGGC CGATTt CAA GAA GAA CAA CAA Ctt
 198> n Tyr Leu Lys Thr II e Gl n Asp Lys Leu Al a Asp Phe Lys Lys Asn Asn Asn Ph
 648 CGT Cgg GAT Cgc CgAT Ct GAGC Ac CgATTt CAA ACC CAA CAA CAC Ct CC T GAC CAA
 216> e Val Gl y II e Al a Asp Leu Ser Thr Asp Tyr Asn His Asn Asn Leu Leu Thr Ly
 702 GTT CCT GAG Cac CGGT AT GGT Ctt CGAAA CCT Ggc CAA GACC GT CC T GAG CAA
 234> s Phe Leu Ser Thr Gl y Met Val Ile Phe Gl u Asn Leu Al a Lys Thr Val Leu Ser As
 756 CCt Gct GgAT GG GAA C Ctg CAA Ggg GAT Gt Gt GAC AT CAG Cca Gc ACC Gt Gt Gt
 252> n Leu Leu Asp Gl y Asn Leu Gl n Gl y Met Leu Asn II e Ser Gl n His Gl n Cys Va
 810 GAA GAA GcA Gt Gt CC Cca GAA CAG Cgg GIG TTT CAG ACA CC T GgAT GAG GAGA
 270> I Lys Lys Gl n Cys Pro Gl n Asn Ser Gl y Cys Phe Arg His Leu Asp Gl u Arg Gl
 864 GgA Gt Gt AA Gigt Ct CC T GAA Ct AC AA GcA GgA AGGT GATA A Gigt Gt GgAAA C
 288> u Gl u Cys Lys Cys Leu Leu Asn Tyr Lys Gl n Gl u Gl y Asp Lys Cys Val Gl u Asn
 919 CC CAAT CCT ACT TGT AAC GA GAA CAAT GGT GGAT GT GAT GC CGATGCC A Gigt ACC G
 307> Pro Asn Pro Thr Cys Asn Gl u Asn Asn Gl y Gl y Cys Asp Al a Asp Al a Lys Cys Thr G
 977 AGGA GgAT TCAGG GAG CAA CGG GAA GAA GAT CAC Cigt Gt Gigt Ac Caa Gcc Tgatt
 326> I u Gl u Asp Ser Gl y Ser Asn Gl y Lys Lys II e Thr Cys Gl u Cys Thr Lys Pro Asp S
 1034 CTT ATCC ACT Gt Gt GAT GGT AT Ctt CTG TAGT
 345> er Tyr Pro Leu Phe Asp Gl y II e Phe Cys Ser

FIG. 1

Applicant(s): Chen et al.

NOVEL MODIFIED MSP-1 NUCLEIC ACID SEQUENCES AND METHODS FOR INCREASING mRNA LEVELS AND PROTEIN EXPRESSION IN CELL SYSTEMS

1 GCAGTAACCTCCGTAATTGATAACATACTTTCTAAAATTGAAATGAATA
 1►AlaValThrProSerValIleAspAsnIleLeuSerLysIleGluAsnGluTyrG
 EcoNI (73)
 56 AGGTTTTATATTTAACCTTAGCAGGTGTTATAGAAGTTAAAAAAACAATT
 19►IuValLeuTyrLeuLysProLeuAlaGlyValTyrArgSerLeuLysGlnLe
 11 AGAAAATAACGTTATGACATTAAATGTTAATGTTAAGGATATTAAATTCACTGA
 37►uGluAsnAsnValMetThrPheAsnValAsnValLysAspIleLeuAsnSerArg
 66 TTTAATAAACGTGAAAATTCAAAAATGTTTAGAATCAGATTAAATTCCATATA
 56►PheAsnLysArgGluAsnPheLysAsnValLeuGluSerAspLeuIleProTyrL
 21 AAGATTTAACATCAAGTAATTATGTTGCAAAGATCCATATAAATTCTTAATAA
 74►ysAspLeuThrSerSerAsnTyrValValLysAspProTyrLysPheLeuAsnLy
 76 AGAAAAAAAGAGATAAATTCTAACGAGTTATAATTATAAGGATICAATAGAT
 92►sGluLysArgAspLysPheLeuSerSerTyrAsnTyrIleLysAspSerIleAsp
 31 ACGGATATAAATTGCAAATGATGTTCTGGATATTATAAAATATTATCCGAAA
 11►ThrAspIleAsnPheAlaAsnAspValLeuGlyTyrTyrLysIleLeuSerGluL
 86 AATATAAAATCAGATTAGATTCAATTAAAAATATCACAGACAAACAAGGTGA
 29►ysTyrLysSerAspLeuAspSerIleLysLysTyrIleAsnAspLysGlnGlyGluL
 41 AAAATGAGAAAATACCTCCCTTTAAACAATATTGAGACCTTATATAAAACAGTT
 47►uAsnGluLysTyrLeuProPheLeuAsnAsnIleGluThrLeuTyrLysThrVal
 196 AATGATAAAAATGATTATTGTAATTCAATTAGAAGCAAAAGTTCTAAATTATA
 66►AsnAspLysIleAspLeuPheValIleHisLeuGluAlaLysValLeuAsnTyrT
 551 CATATGAGAAAATCAAACGTAGAAGTTAAAATAAAAGAACTTAATTACTTAAAC
 184►hrTyrGluLysSerAsnValGluValLysIleLysGluLeuAsnTyrLeuLysTh
 506 AATTCAAGACAAATTGGCAGATTAAAAAAATAACAATTCTGTGGAATTGCT
 202►rIleGlnAspLysLeuAlaAspPheLysLysAsnAsnAsnPheValGlyIleAla
 661 GATTATCAACAGATTATAACCATAATAACTTATTGACAAAGTTCTTAGTACAG
 221►AspLeuSerThrAspTyrAsnHisAsnAsnLeuLeuThrLysPheLeuSerThrG
 716 GTATGGTTTGAAATCTTGCTAAACCGTTTATCTAATTACTTGATGGAAA
 239►IyMetValPheGluAsnLeuAlaLysThrValLeuSerAsnLeuLeuAspGlyAs
 771 CTTGCAAGGTATGTTAACATTTCACAAACCCAATGCGTAAAAAAACAATGTCCA
 257►nLeuGlnGlyMetLeuAsnIleSerGlnHisGlnCysValLysGlnCysPro
 826 CAAAATTCTGGATTTTCAGACATTAGATGAAAGAGAAGAATGTAATGTTTAT
 276►GlnAsnSerGlyCysPheArgHisLeuAspGluArgGluGluCysLysCysLeuL
 881 TAAATTACAAACAAGAAGGTGATAAAATGTGTTGAAAATCCAAATCCTACTTGTAA
 294►euAsnTyrLysGlnGluGlyAspLysCysValGluAsnProAsnProThrCysAs
 936 CGAAAATAATGGTGGATGTGATGCCAATGTACCGAAGAAGATTCAAGGT
 312►nGluAsnAsnGlyGlyCysAspAlaAspAlaLysCysThrGluGluAspSerGly
 991 AGCAACGGAAAGAAAATCACATGTGAAATGTACTAAACCTGATTCTTATCCACTTT
 331►SerAsnGlyLysLysIleThrCysGluCysThrLysProAspSerTyrProLeuP

PstI (1059)

1046 TCGATGGTATTTCTGCAGTCACCACCAACCACCAACT
349 ► heAspGl y l ePheCysSer Hi s Hi s Hi s Hi s Hi s • • •

FIG. 2

NOVEL MODIFIED MSP-1 NUCLEIC ACID SEQUENCES AND
METHODS FOR INCREASING mRNA LEVELS AND PROTEIN
EXPRESSION IN CELL SYSTEMS

Codon	AA	goat b-casein	goat K-casein	MSP wt	Edited MSP	mouse b-casein	mouse a-casein	mouse g-casein	mouse e-casein
TTT	Phe	5	4	8	0	4	8	3	4
TTC	Phe	4	3	7	15	4	6	7	1
TTA	Leu	0	2	25	0	0	0	0	0
TTG	Leu	0	2	3	0	0	0	0	1
TCT	Ser	5	1	4	1	13	5	7	5
TCC	Ser	2	2	2	3	6	14	8	2
TCA	Ser	1	4	10	1	1	3	2	0
TCG	Ser	0	1	0	0	0	0	0	0
TAT	Tyr	2	7	17	2	1	3	2	1
TAC	Tyr	1	2	3	18	2	6	6	7
TAA	***	1	2	0	0	1	0	1	0
TAG	***	0	0	0	0	0	0	0	0
TGT	Cys	1	1	10	12	0	0	1	0
TGC	Cys	0	2	2	0	2	2	2	1
TGA	***	0	0	0	0	0	1	0	1
TGG	Trp	1	1	0	0	0	2	2	2
CTT	Leu	9	1	9	0	16	9	3	3
CTC	Leu	5	2	0	0	7	8	0	1
CTA	Leu	1	2	1	0	1	2	1	0
CTG	Leu	11	5	0	38	10	17	4	1
CCT	Pro	17	6	4	2	8	6	3	0
CCC	Pro	12	0	1	6	8	6	6	4
CCA	Pro	3	13	5	1	5	6	2	2
CCG	Pro	1	1	0	1	0	0	0	1
CAT	His	0	1	3	0	2	6	2	1
CAC	His	5	3	1	4	4	0	3	0
CAA	Gln	5	9	9	0	9	21	9	7
CAG	Gln	16	6	0	9	21	32	12	8
CGT	Arg	0	1	1	0	0	0	0	0
CGC	Arg	0	0	0	0	1	0	0	0
CGA	Arg	0	0	1	0	0	0	0	0
CGG	Arg	1	0	0	3	0	0	0	0
ATT	Ile	4	5	13	0	3	4	3	4
ATC	Ile	6	3	2	20	7	5	8	5
ATA	Ile	1	3	5	0	1	0	2	0
ATG	Met	7	3	3	3	4	12	2	13
ACT	Thr	7	6	3	2	6	5	1	4
ACC	Thr	2	7	3	13	4	4	4	4
ACA	Thr	2	4	9	1	1	1	2	0
ACG	Thr	0	0	1	0	0	0	2	0
AAT	Asn	2	6	29	3	4	6	3	1
AAC	Asn	2	3	12	38	4	9	4	6
AAA	Lys	7	6	38	0	6	7	3	5
AAG	Lys	6	4	4	42	3	6	13	7
AGT	Ser	2	6	5	2	3	6	6	5
AGC	Ser	5	0	2	16	2	6	6	3
AGA	Arg	2	2	4	3	1	8	1	1
AGG	Arg	0	2	0	0	0	0	0	1
GTT	Val	5	6	15	0	7	4	2	3
GTC	Val	8	2	1	11	7	3	3	0
GTA	Val	2	2	5	0	2	4	1	3
GTG	Val	8	4	0	10	6	3	5	3
GCT	Ala	1	3	2	0	8	17	4	2
GCC	Ala	4	7	1	8	6	3	3	3
GCA	Ala	3	7	6	1	4	13	1	1
GCG	Ala	0	1	0	0	0	0	0	0
GAT	Asp	4	5	25	27	3	6	4	2
GAC	Asp	0	2	2	0	1	2	1	3
GAA	Glu	10	6	21	3	6	12	9	6
GAG	Glu	9	5	4	22	5	5	5	5
GGT	Gly	2	1	8	4	0	0	0	0
GGC	Gly	0	0	0	0	0	0	0	0
GGA	Gly	2	1	6	3	1	0	1	0
GGG	Gly	1	0	0	7	1	0	0	0

FIG. 3A

NOVEL MODIFIED MSP-1 NUCLEIC ACID SEQUENCES AND
METHODS FOR INCREASING mRNA LEVELS AND PROTEIN
EXPRESSION IN CELL SYSTEMS

Codon	AA	MSP wt	Edited MSP	MSP wt	Edited MSP	E.coli	Human
TTT	Phe	8	0	0.53	0	0.5	0.35
TTC	Phe	7	15	0.47	1	0.5	0.65
TTA	Leu	25	0	0.66	0	0.11	0.05
TTG	Leu	3	0	0.08	0	0.11	0.09
TCT	Ser	4	1	0.17	0.04	0.27	0.17
TCC	Ser	2	3	0.09	0.13	0.21	0.26
TCA	Ser	10	1	0.43	0.04	0.13	0.11
TCG	Ser	0	0	0	0	0.14	0.07
TAT	Tyr	17	2	0.85	0.1	0.54	0.47
TAC	Tyr	3	18	0.15	0.9	0.46	0.53
TAA	***	0	0				
TAG	***	0	0				
TGT	Cys	10	12	0.83	1	0.45	0.3
TGC	Cys	2	0	0.17	0	0.55	0.7
TGA	***	0	0				
TGG	Trp	0	0	0	0	1	1
CTT	Leu	9	0	0.24	0	0.12	0.11
CTC	Leu	0	0	0	0	0.12	0.22
CTA	Leu	1	0	0.02	0	0.03	0.07
CTG	Leu	0	38	0	1	0.72	0.46
CCT	Pro	4	2	0.4	0.2	0.14	0.24
CCC	Pro	1	6	0.1	0.6	0.11	0.41
CCA	Pro	5	1	0.5	0.1	0.2	0.24
CCG	Pro	0	1	0	0.1	0.54	0.11
CAT	His	3	0	0.75	0	0.64	0.42
CAC	His	1	4	0.25	1	0.36	0.58
CAA	Gln	9	0	1	0	0.31	0.26
CAG	Gln	0	9	0	1	0.69	0.74
CGT	Arg	1	0	0.17	0	0.46	0.09
CGC	Arg	0	0	0	0	0.32	0.19
CGA	Arg	1	0	0.17	0	0.05	0.1
CGG	Arg	0	3	0	0.5	0.06	0.15
ATT	Ile	13	0	0.65	0	0.39	0.23
ATC	Ile	2	20	0.1	1	0.52	0.64
ATA	Ile	5	0	0.25	0	0.08	0.13
ATG	Met	3	3	1	1	1	1
ACT	Thr	3	2	0.19	0.13	0.36	0.2
ACC	Thr	3	13	0.19	0.81	0.38	0.47
ACA	Thr	9	1	0.56	0.06	0.09	0.21
ACG	Thr	1	0	0.06	0	0.17	0.12
AAT	Asn	29	3	0.71	0.07	0.29	0.34
AAC	Asn	12	38	0.29	0.93	0.71	0.66
AAA	Lys	38	0	0.9	0	0.72	0.45
AAG	Lys	4	42	0.1	1	0.28	0.55
AGT	Ser	5	2	0.21	0.09	0.11	0.11
AGC	Ser	2	16	0.09	0.7	0.14	0.29
AGA	Arg	4	3	0.67	0.5	0.08	0.24
AGG	Arg	0	0	0	0	0.03	0.23
GTT	Val	15	0	0.71	0	0.37	0.13
GTC	Val	1	11	0.05	0.52	0.12	0.27
GTA	Val	5	0	0.24	0	0.28	0.09
GTG	Val	0	10	0	0.48	0.23	0.5
GCT	Ala	2	0	0.22	0	0.33	0.31
GCC	Ala	1	8	0.11	0.89	0.18	0.4
GCA	Ala	6	1	0.67	0.11	0.28	0.17
GCG	Ala	0	0	0	0	0.21	0.12
GAT	Asp	25	27	0.93	1	0.48	0.38
GAC	Asp	2	0	0.07	0	0.52	0.62
GAA	Glu	21	3	0.84	0.12	0.67	0.4
GAG	Glu	4	22	0.16	0.88	0.33	0.6
GGT	Gly	8	4	0.57	0.29	0.46	0.15
GGC	Gly	0	0	0	0	0.4	0.44
GGA	Gly	6	3	0.43	0.21	0.06	0.17
GGG	Gly	0	7	0	0.5	0.08	0.24

FIG. 3B

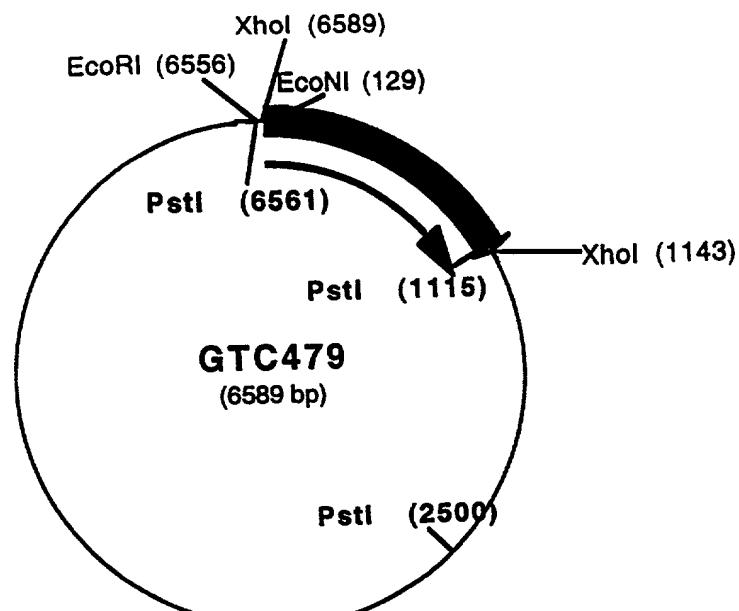


FIG. 4A

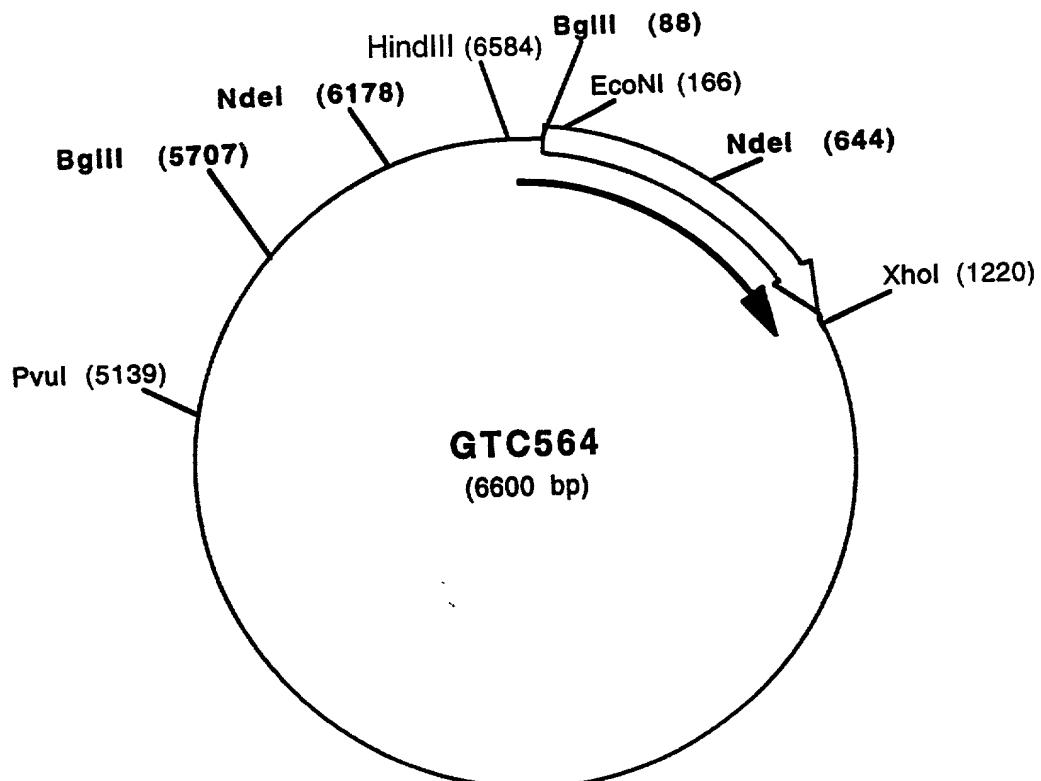


FIG. 4B

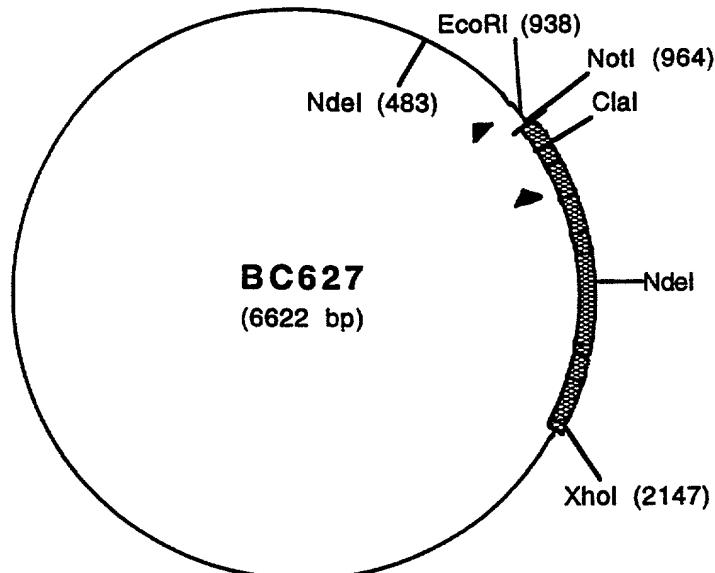


FIG. 4C

Oligos used:

OT1:

TCG ACG AGA GCC ATG AAG GTC CTC ATC CTT GCC TGT CTG GTG GCT
CTG GCC ATT GCA AGA GAG CAG GAA GAA CTC AAT GTA GTC GGT A,

OT2:

GAT CTA CCG ACT ACA TTG AGT TCT TCC TGC TCT CTT GCA ATG GCC
AGA GCC ACC AGA CAG GCA AGG ATG AGG ACC TTC ATG GCT CTC G,

MSP1:

AATAGATCTCGAGTAACTCCTCCGTAATTG,

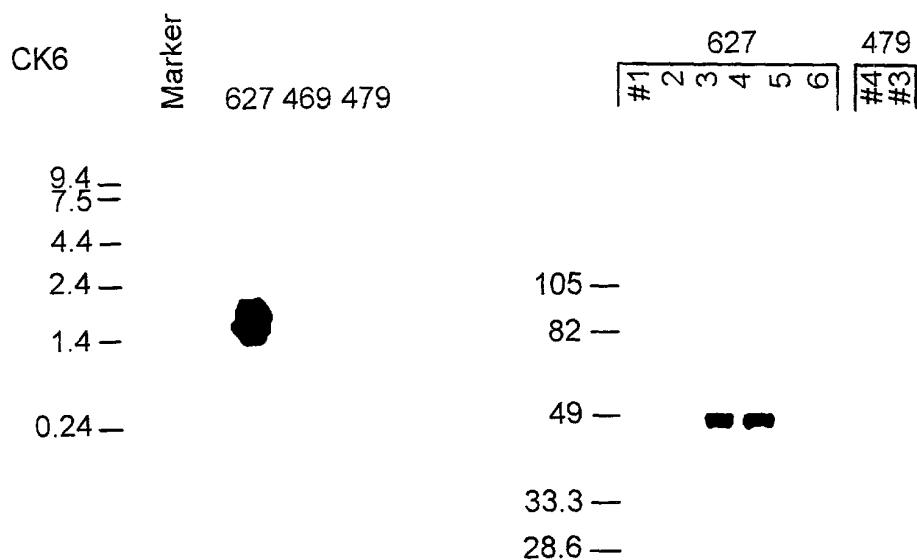
MSP2:

AATTCTCGAGTTAGTGGTGGTGGTGGTGGTACTGCAGAAATACCATC

MSP8:

TAACTCGAGCGAACCATGAAGGT CCTCATCCTGCCTGTCTGGTGGCTCTGG
CCATTGCA

FIG. 6



PANEL A

PANEL B

FIG. 5

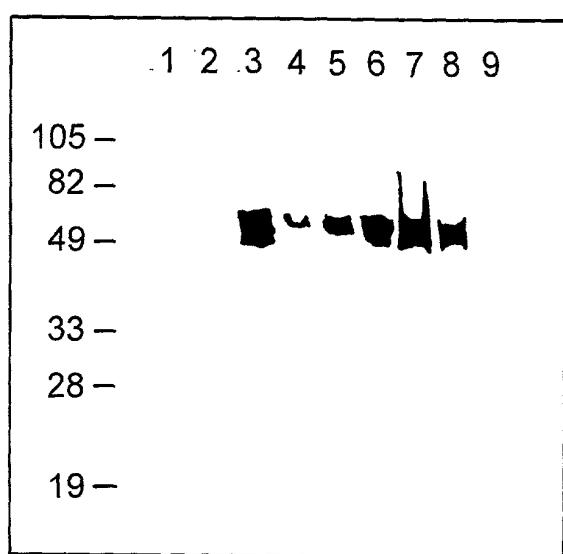


FIG. 10

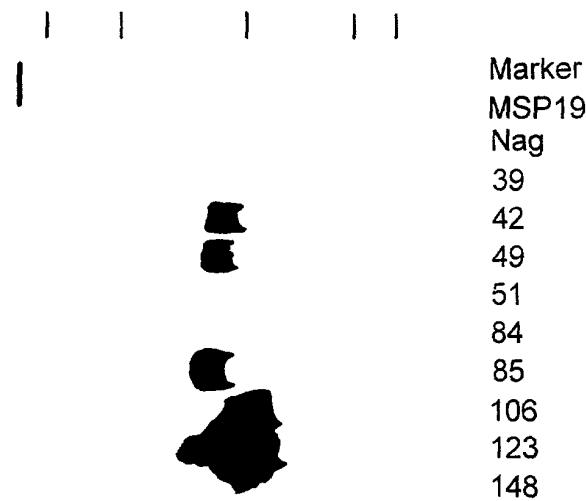


FIG. 13

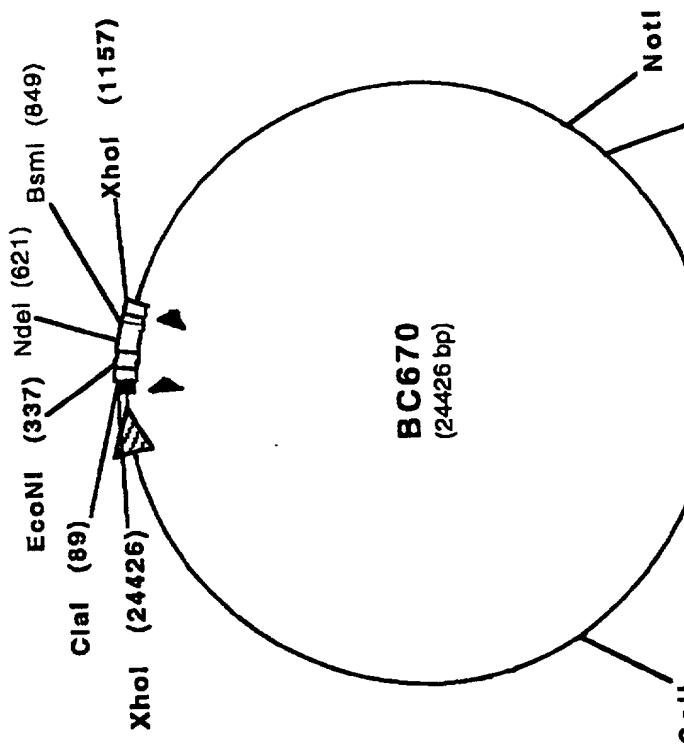


FIG. 7

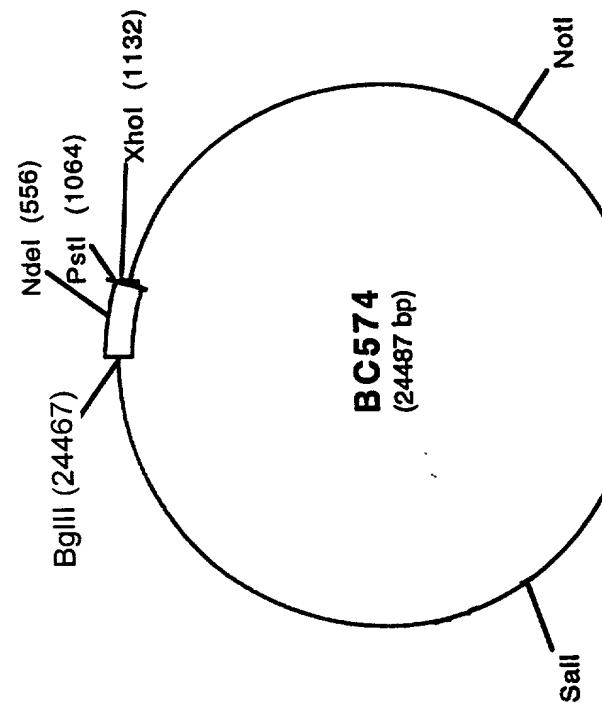


FIG. 8

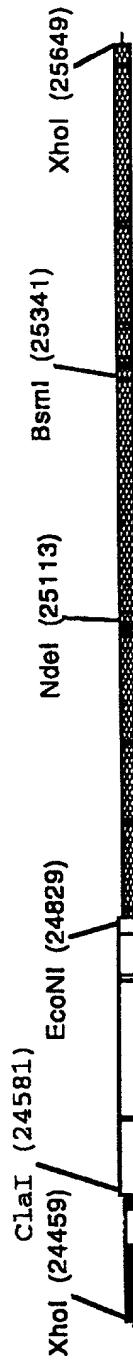


Diagram of BC620

FIG. 9

NOVEL MODIFIED MSP-1 NUCLEIC ACID SEQUENCES AND
METHODS FOR INCREASING mRNA LEVELS AND PROTEIN
EXPRESSION IN CELL SYSTEMS

26 ATGAAGGTCCCTCATATAATTGCCCTGCTGGTGGCTCTGGCATTGCAGCGTCACTCCCTCCGTATCGATAAC

1► M K V L I I A C L V A L A I A A V T P S V I D N
98 ATCCCTGTCCAAGATCGAGAACGGAGTACGAGGTGCTGTACCTGAAGGCCCTGGCAGGAGTCTACAGGAGCT

25► I L S K I E N E Y E V L Y L K P L A G V Y R S L
169 GAAGAAGCAGCTGGAGAACAACTGTATGACCTCAACGTGAAGGATATCCTGAACAGCAAGGTTCAA

48► K K Q L E N N V M T F N V N V K D I L N S R F N
241 CAAGAAGGGAGAACTTCAAGAACGTGCTGGAGAGCGATCTGATCCCTACAAGGATCTGACCAGCAGCAACTA

72► K R E N F K N V L E S D L I P Y K D L T S S N Y
EcoNI (337)

313 CGGGTCAAAGATCCCTACAAGTCTGAAACAAGGAGAAGAGAGATAAGTCTGAGCAGTACAATTACAT →

96► V V K D P Y K F L N K E K R D K F L S S Y N Y I
385 CAAGGATAGCATTGACACCGATATCAACTTCGCCAACGATGTCTGGATACTACAAGATCTGTCGGAGAA

120► K D S I D T D I N F A N D V L G Y Y K I L S E K
457 GTACAAAGAGCGATCTGGATAGCATCAAGAACGTACATCAACGATAAGCAGGGAGAGAACGAGAAGTACCTGCC

144► Y K S D L D S I K K Y I N D K Q G E N E K Y L P
529 CTTCTGAACAACATCGAGACCCCTGTACAAGAACCGTCAACGATAAGATTGATCTGTCGATCCACCTGGA

168► F L N . N I E T L Y K T V N D K I D L F V I H L E
NdeI (621)

601 GGCAAGGTCTGCAGTACACATATGAGAAGAGCAACGTGGAGGTCAAGATCAAGGAGCTGAATTACCTGAA ←

192► A K V L Q Y T Y E K S N V E V K I K E L N Y L K
673 GACCATCCAGGATAAGCTGGCGATTCAAGAACAAACTTCGTCGQAATGCCGATCTGAGCACCGA

216► T I Q D K L A D F K K N N N F V G I A D L S T D
745 TTACAACCACAAACAAACCTGCTGACCAAGTCTGAGCACCGGAATGGTCCTGAAAACCTGCCAACCGT

240► Y N H N N L L T K F L S T G M V F E N L A K T V
BsmI (849)

817 CCTGAGCAACCTGGTGGATGQAAACCTGCAGGGATGCTGCAAGATCAGCCAGCACCAAGTGTGAAGAAC →

264► L S N L L D G N L Q G M L Q I S Q H Q C V K K
888 AGTGTCCCCAGAACAGCGGATGCTCAGACACCTGGATGAGAGGGAGGTGCAAGTGCCTGCTGAACTA

288► Q C P Q N S G C F R H L D E R E E C K C L L N Y
958 CAAGCAGGAAGGAGATAAGTGTGTGAAAACCCAACTCTACTTGTAACGAGAACAAATGGAGGATGCGATG

311► K Q E G D K C V E N P N P T C N E N N G G C D
1029 CCGATGCCAAGTGTACCGAGGAGGATTCAAGAAGCAACGGAAAGAAGATCACCTGCGAGTGTACCAAGCCT

335► A D A K C T E E D S G S N G K K I T C E C T K P
XbaI (1157)

1100 GATTCTTATCCACTGTCGATGGtatTTCTGCAGTCACCACCACCAACTCGAGGAT ←

359► D S Y P L F D G I F C S H H H H H H • L E D

FIG. 11

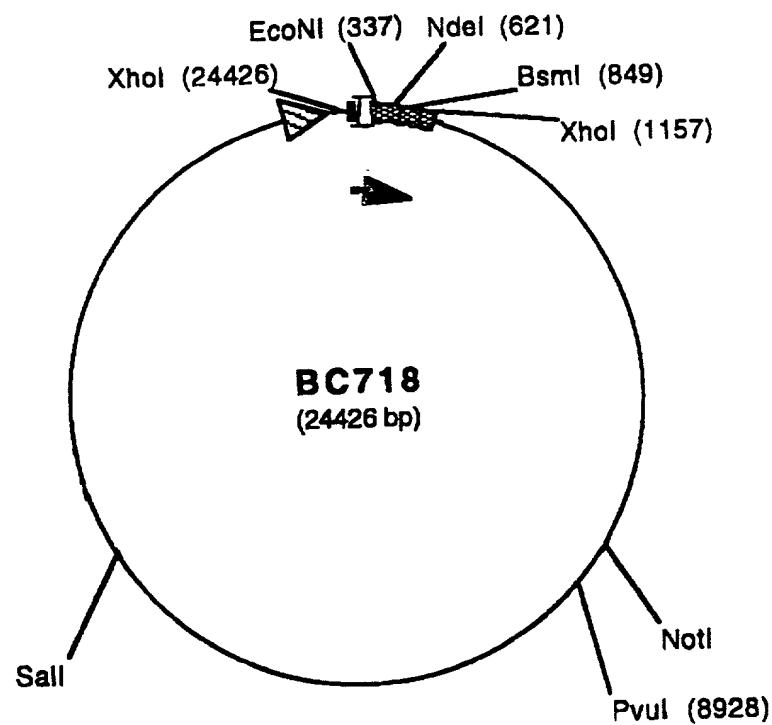


FIG. 12